

WHAT IS CLAIMED IS:

1. A web site system, comprising:

a web server system that is responsive to requests from online users by generating and returning web pages, wherein the web server system includes one or more applications the generate personalized content for recognized users based on browse histories of such users; and

an event history server that persistently stores event data descriptive of events that occur during browsing sessions of each of a plurality of users of the web server system, and makes such event data available in real time to the one or more applications to facilitate personalization of web pages for the users;

wherein the event history server implements a query interface through which the one or more applications can retrieve the event data of a given user by event type and by event time of occurrence.

2. The web site system of Claim 1, wherein the event history server records the event data for a given event as an event object that includes at least the following: an event type identifier, an event value, a user ID, and a time stamp.

3. The web site system of Claim 1, wherein the event history server includes at least one storage layer server that stores the event data persistently by user ID, and further includes at least one cache layer server that caches event data of online users.

4. The web site system of Claim 2, wherein the cache layer server is configured to collect event data of an unrecognized user during a browsing session, and to pass such collected event data to the at least one storage layer server for persistent storage thereof if the unrecognized user becomes recognized during the browsing session.

5. The web site system of Claim 1, wherein the event history server comprises a plurality of cache layer servers, each of which is assigned to a different respective set of browse session IDs such that a given user remains assigned to a particular cache layer server throughout a browse session.

6. The web site system of Claim 1, wherein the event history server comprises a plurality of mirrored storage layer servers that persistently store like event data by user ID.

7. The web site system of Claim 1, wherein the query interface of the event history server supports queries of the form “has User X accessed URL Y?”

8. The web site system of Claim 1, wherein the query interface of the event history server supports queries of the form “when has User X accessed URL Y?”

9. The web site system of Claim 1, wherein the event history server records event data for substantially every mouse click action of every recognized user of a corresponding web site.

10. The web site system of Claim 1, wherein the event history server records impression event data indicative of specific items presented to users on dynamically generated web pages.

11. The web site system of Claim 1, wherein the at least one application includes a web search application that provides functionality for searching an index of web pages, and uses the event history server to identify and highlight web search result items that have previously been accessed by a user conducting a current search.

12. The web site system of Claim 1, wherein the at least one application includes an application that provides functionality for users to interactively view and organize their respective browse history data as recorded by the event history server.

13. The web site system of Claim 1, wherein the event history server generates user-specific Bloom filters reflective of event histories of specific users, and uses the user-specific Bloom filters to respond to queries from the at least one application.

14. An event history server, comprising:

a plurality of storage layer servers that persistently store, and provide real-time access to, event data descriptive of browsing events that occur during web browsing sessions of users, wherein the storage layer servers store the event data in an indexed form in association with corresponding user IDs, and are implemented as mirrors of each other such that client requests for persistently stored event data can be serviced by any one of the storage layer servers; and

a plurality of cache layer servers that serve as intermediaries between the storage layer servers and clients of the event history server, wherein each cache layer

server stores cached event data within a respective cache and uses the cached event data to respond to queries from the clients.

15. The event history server of Claim 14, wherein browsing sessions are uniquely assigned to specific cache layer servers such that each cache layer server caches event data for a different respective set of users.

16. The event history server of Claim 14, wherein the cache layer servers are responsive to updates received from an event reporting component by updating their respective caches with event data specified by such updates, and by forwarding such updates to the storage layer servers for persistent storage of the event data specified therein.

17. The event history server of Claim 14, wherein the storage layer servers store the event data for a given event as an event object that includes at least the following: an event type identifier, and event value, and a time stamp.

18. The event history server of Claim 17, wherein the storage layer servers and the cache layer servers implement a query interface through which clients of the event history server can retrieve event objects based on at least the following criteria: user ID, event time, event date.

19. The event history server of Claim 14, wherein each cache layer server is configured to collect event data of an unrecognized user during a browsing session, and to pass such collected event data to a storage layer server for persistent storage thereof if the unrecognized user becomes recognized during the browsing session.

20. The event history server of Claim 14, wherein the cache layer servers support queries of the form "has User X accessed URL Y?"

21. The event history server of Claim 14, wherein the cache layer servers support queries of the form "when has User X accessed URL Y?"

22. The event history server of Claim 14, wherein the storage layer servers store event data for substantially every mouse click action of every recognized user of a corresponding web site.

23. The event history server of Claim 14, wherein the storage layer servers record impression event data indicative of specific items presented to users on dynamically generated web pages.

24. A method of processing a search query from a user, comprising, on a server system that is remote from the user:

executing the search query to generate a query result, wherein the query result comprises URLs of web pages that are responsive to the search query;

for at least one of the URLs, querying a server to determine whether event data stored for said user indicates that the user previously accessed the URL; and

if the event data indicates that the user previously accessed the URL, including within a search results page an annotation indicating to the user that the URL was previously accessed.

25. The method of Claim 24, wherein the annotation further indicates a date that the user accessed the URL, as indicated by the event data stored for the user.

26. The method of Claim 24, wherein the server responds to the querying at least in part by using a Bloom filter stored for the user to evaluate whether the user previously accessed the URL.

27. The method of Claim 24, wherein the method further comprises querying said server to identify a date that the user accessed the URL, and including said date within the search results page.

28. The method of Claim 24, wherein the method comprises querying the server separately for each of a plurality of said URLs to determine whether each such URL was previously accessed by the user, as indicated by the event data stored for the user.

29. A search engine system comprising executable query processing code that embodies the method of Claim 24.

30. A system that provides functionality for conducting Internet searches, the system comprising:

a search application that is responsive to search queries from users by generating and returning search results pages listing search result URLs of external web pages that are responsive to such search queries; and

a server that records user-specific data indicative of the search result URLs selected by users during browsing of the search result pages;

wherein the search application accesses the server to determine whether specific search result URLs have previously been accessed by a user who is conducting a search, and incorporates into the search results pages indications of which search result URLs were previously accessed, whereby users are notified of search result URLs they have previously accessed.

31. The system of Claim 30, wherein the server additionally records information indicative of the dates that specific search result URLs were accessed by specific users, and the search application uses said information to further indicate within the search results pages the dates that specific search result URLs were accessed.

32. The system of Claim 30, wherein the server is configured to generate user-specific Bloom filters that reflect search result URLs selected by users, and to use said user-specific Bloom filters to evaluate whether specific users have previously accessed specific URLs.

33. The system of Claim 30, wherein the system facilitates tracking of a user's selection of a search result URL by redirecting the user to the search result URL in response to selection thereof.

34. The system of Claim 30, wherein the server further records user-specific data indicative of specific search queries submitted by specific users, and the search application uses the data stored by the server to identify and highlight search result URLs that did not come up when a user conducting a current search previously conducted the same search.

35. The system of Claim 30, wherein the server persistently stores event data for each of a plurality of event types, and implements a query set that provides for real-time retrieval of user-specific event data by event type.

36. A method of providing browse-history-based personalization of search results, the method comprising:

maintaining event history data indicative of search result URLs selected by a user during browsing of search results pages;

generating a user-specific Bloom filter that reflects a plurality of the search result URLs selected by the user as indicated within said event history data;

in response to an occurrence of a given URL within a result of a search query submitted by the user, determining whether the user previously accessed the given URL at least in part by analyzing the user-specific Bloom filter; and

when the user is determined to have previously accessed the given URL, personalizing a search results page for the user with an indication that the user previously accessed the given URL.

37. The method of Claim 36, further comprising maintaining a record of when each such search result URL was selected by the user, and personalizing the search results page with an indication of when the user accessed the given URL.

38. The method of Claim 36, wherein the user-specific Bloom filter reflects accesses by the user to URLs of a plurality of independent web sites.

39. The method of Claim 36, wherein the user-specific Bloom filter reflects actions performed by the user over multiple browsing sessions.

40. The method of Claim 36, wherein the step of determining whether the user previously accessed the given URL comprises:

initially determining whether the Bloom filter indicates a possible access by the user to the given URL; and

when the Bloom filter indicates such a possible access, accessing said event history data to determine whether an access to the given URL is in fact reflected therein.

41. The method of Claim 36, wherein the step of determining whether the user previously accessed the given URL comprises:

initially determining whether the Bloom filter indicates a possible access by the user to the given URL; and

when the Bloom filter indicates such a possible access, treating the given URL as having previously been accessed by the user, such that the user's event history data need not be directly accessed.

42. A method of providing browse-history-based content personalization, the method comprising:

- maintaining a browse history describing a user's interactions with a web site;
- generating a user-specific Bloom filter that represents a plurality of events included within the user's browse history;
- determining whether a particular event exists within the browse history of the user at least in-part by analyzing said Bloom filter; and
- personalizing a web page for the user such that the web page reflects whether the particular event exists within the browse history of the user.

43. The method of Claim 42, wherein the Bloom filter reflects accesses by the user to search result URLs, and the step of determining whether a particular event exists within the browse history comprises using the Bloom filter to evaluate whether the user previously accessed a particular search result URL.

44. The method of Claim 42, wherein the Bloom filter reflects a plurality of items selected by the user from an electronic catalog, and wherein the step of determining whether a particular event exists within the browse history of the user comprises using the Bloom filter to evaluate whether the user previously selected a particular item from the electronic catalog.

45. The method of Claim 42, wherein the Bloom filter reflects events in which an item was merely displayed to but not selected by the user, and wherein determining whether a particular event exists within the browse history of the user comprises using the Bloom filter to evaluate whether a particular item was previously displayed to the user.